Aorsin, a novel serine proteinase with trypsin-like specificity at acidic pH

Byung Rho LEE*[1], Masato FURUKAWA†, Koichiro YAMASHITA*, Yurie KANASUGI*, Choko KAWABATA‡, Kenichi HIRANO§, Kenichi ANDO§ and Eiji ICHISHIMA*†

*Laboratory of Molecular Enzymology, Department of Bioengineering, Faculty of Engineering, Soka University, Hachioji, Tokyo 192-8577, Japan, †Laboratory of Molecular Enzymology, Division of Bioengineering, Graduate School of Engineering, Soka University, Hachioji, Tokyo 192-8577, Japan, ‡Technical Research Center of T. Hasegawa Co., Ltd., Kariyado, 335 Nakahara-ku, Kawasaki 211-0022, Japan, and §Gifu R and D Center of Amano Enzyme Inc., 4-179-35, Sue-cho, Kakamigahara, Gifu 509-0108, Japan

A proteinase that hydrolyses clupeine and salmine at acidic pH, called aorsin, was found in the fungus *Aspergillus oryzae*. Purified aorsin also hydrolysed benzyloxycarbonyl-Arg-Arg-4-methylcoumaryl-7-amide optimally at pH 4.0. The specificity of aorsin appeared to require a basic residue at the P_1 position and to prefer paired basic residues. Aorsin activated plasminogen and converted trypsinogen to trypsin. The trypsin-like activity was inhibited strongly by antipain or leupeptin, but was not inhibited by any other standard inhibitors of peptidases. To identify the catalytic residues of aorsin, a gene was cloned and an expression system was established. The predicted mature protein of aorsin was 35% identical with the classical late-infantile neuronal ceroid lipofuscinosis protein CLN2p and was 24% identical with *Pseudomonas* serine-carboxyl proteinase, both of which are pepstatin-insensitive carboxyl proteinases. Several putative catalytic residues were mutated. The k_{cat}/K_m values of the mutant enzymes $Glu^{86} \rightarrow Gln$, $Asp^{211} \rightarrow Asn$ and $Ser^{354} \rightarrow Thr$ were 3–4 orders of magnitude lower and $Asp^{90} \rightarrow Asn$ was 21-fold lower than that of wild-type aorsin, indicating that the positions are important for catalysis. Aorsin is another of the S53 family serine-carboxyl proteinases that are not inhibited by pepstatin.

Key words: acid protease, aorsin, *Aspergillus oryzae*, pepstatin-insensitive carboxyl proteinase, serine proteinase, trypsin-like specificity.

INTRODUCTION

Acid proteinases that are active at acidic pH have been referred to as aspartic proteinases [1]. Generally, members of the aspartic proteinase family are composed of two homologous domains, each containing a catalytic aspartic acid residue in a consensus sequence of Asp-Thr/Ser-Gly, in close proximity, to allow their catalytic function [1]. These enzymes are inhibited specifically by pepstatin [2].

In addition to these aspartic proteinases, there exist some other distinct families that do not contain the consensus sequence of Asp-Thr/Ser-Gly. These are called non-pepsin-type aspartic proteinases and include aspergillopepsin II (EC 3.4.23.19) [3], scytalidopepsin B (EC 3.4.23.32) [4], pseudomonapepsin (EC 3.4.23.37) [5,6] and xanthomonapepsin (EC 3.4.23.33) [7]. These families of aspartic proteinases are not sensitive to pepstatin, hence they are also called pepstatin-insensitive carboxyl proteinases.

However, another important member of the aspartic proteinases (of family A7) is the classical late-infantile neuronal ceroid lipofuscinosis (LINCL) protein, CLN2p (ceroid lipofuscinosis, neuronal 2 protein), a human lysosomal enzyme that, when mutated, leads to the fatal neurodegenerative disease, LINCL [8–10]. CLN2p has both tripeptidyl-peptidase I and endopeptidase activity at acidic pH [11,12]. The sequence of CLN2p is significantly similar to the two bacterial pepstatin-insensitive carboxyl proteinases, pseudomonapepsin and xanthomonapepsin [6,8,13]. Within the sequences of these enzymes, Rawlings and Barrett [11] identified a significant conserved sequence, Gly-Xaa-Ser, which is characteristic of the active-site motif of many serine peptidases. More recently, Lin et al. [14] have shown that Ser^{475}, Asp^{360} and Asp^{517} are essential for the activity of CLN2p/tripeptidyl-peptidase I, and also that a di-isopropyl fluorophosphate specifically reacts with CLN2p at Ser^{475}. Furthermore, the crystal structure of pseudomonapepsin has been determined [15]. On the basis of its three-dimensional structure, Wlodawer et al. [15] have proposed that a family containing pseudomonapepsin can be defined as a novel family of subtilisin-like enzymes that have a unique catalytic triad, Glu^{80}, Asp^{84} and Ser^{287}, and that the enzymes in this family are called serine-carboxyl proteinases [pseudomonapepsin is thus also known as *Pseudomonas* serine-carboxyl proteinase (PSCP)]. This new family was classified as S53 in the MEROPS database [15a]. Consequently, the EC numbers for pseudomonapepsin and xanthomonapepsin have been reassigned as EC 3.4.21.100 and 3.4.21.101 respectively.

In general, acid proteinases prefer hydrophobic or bulky amino acid residues at the P_1 position of the substrate. But in the present study, we found a novel acid proteinase, which we called aorsin, and which prefer basic residues at acidic pH. The unique specificity of aorsin is described here. To define its catalytic residues, we carried out mutational experiments on aorsin whereby mutated enzymes were constructed and characterized.

Abbreviations used: Boc, t-butoxycarbonyl; CLN2p, ceroid lipofuscinosis, neuronal 2 protein; KSCP, kumamolysin serine-carboxyl proteinase; LINCL, classical late-infantile neuronal ceroid lipofuscinosis; MCA, 4-methylcoumaryl-7-amide; niaD300, nitrate-assimilation-defective; PSCP, *Pseudomonas* serine-carboxyl proteinase; $t_{1/2}$, half-life; T_m, melting temperature; XCP, *Xanthomonas* pepstatin-insensitive carboxyl proteinase; Z, benzyloxycarbonyl.

[1] To whom correspondence should be addressed (e-mail brlee@t.soka.ac.jp).

The nucleotide sequence data of the aorsin *aorO* gene from *Aspergillus oryzae* appear in the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the Accession Number AB084899.

EXPERIMENTAL

Materials and reagents

Protease M powder was supplied by Amano Enzyme Inc. (Nagoya, Japan), and tyrostatin was supplied by Professor K. Oda. All fluorogenic peptide substrates were purchased from the Peptide Institute (Osaka, Japan). Plasminogen and trypsinogen were purchased from Roche Diagnostics (Mannheim, Germany) and Sigma Chemical Co. (St. Louis, MO, U.S.A.) respectively. *Flavobacterium* N-glycosidase F was purchased from Roche. The restriction endonucleases were purchased from Takara Shuzo (Kyoto, Japan). All other chemicals were of reagent grade and were purchased from various commercial firms.

Strains, plasmids and media

Aspergillus oryzae RIB40 was used as the source of genomic DNA. A nitrate-assimilation-defective mutant of *A. oryzae* niaD300 (*niaD*$^-$) was used as the recipient strain for transformation [16]. A Czapec-Dox plate containing 0.3 % (w/v) $NaNO_3$, 0.2 % (w/v) KCl, 0.1 % (w/v) KH_2PO_4, 0.05 % (w/v) $MgSO_4 \cdot 7H_2O$, 0.002 % (w/v) $Fe_2SO_4 \cdot 7H_2O$, 2 % (w/v) glucose and 2 % (w/v) agar was used for the selection of transformants. DPYM medium (pH 5.0) containing 2 % (w/v) dextrin, 1 % (w/v) polypeptone, 0.5 % (w/v) yeast extract, 0.5 % (w/v) KH_2PO_4, 0.05 % (w/v) $MgSO_4 \cdot 7H_2O$ and 50 mM maleic acid was used for enzyme production.

Assay of enzyme activity

The hydrolysis of the fluorogenic peptide substrates was spectrofluorimetrically measured at 30 °C. The reaction mixtures contained 945 μl of 50 mM sodium acetate buffer (pH 4.0), 5 μl of 10 mM substrate solution and 50 μl of sample enzyme in a total volume of 1 ml. The increase in fluorescence intensity produced by substrate cleavage during the incubation was measured at an emission wavelength of 440 nm with excitation at 360 nm using a fluorescence spectrophotometer. One katal of the enzyme activity is defined as the amount of enzyme that liberates 1 mol of 4-methylcounmaryl-7-amide (MCA) from the fluorogenic peptide substrate/s at 30 °C and pH 4.0.

Kinetic analysis

For each kinetic experiment, the assays were performed with at least eight different substrate concentrations, which were prepared at a concentration of one fourth to one half of the K_m. Initial velocities (V) were determined and plotted against substrate concentration [S]. The data were fitted to the Michaelis–Menten equation $V = V_{max} \cdot [S]/(K_m + [S])$ with a nonlinear regression analysis program. The best fits of the data produced V_{max} and K_m values, where V_{max} represents the maximum rate of hydrolysis and K_m is the Michaelis constant. The catalytic-centre activity (k_{cat}) values were calculated from $V_{max} = k_{cat} \cdot [E]$, where [E] represents enzyme concentration. Final values of $k_{cat,}$ K_m and k_{cat}/K_m were determined by the means obtained for three to five repeats.

Purification procedure

Commercial protease M powder was used as the starting material for the purification of the enzyme. All subsequent procedures were carried out at 4 °C. Protease M powder (10 g) was dissolved in 100 ml of 10 mM acetate buffer, pH 5.0 (buffer A), 40 g of solid ammonium sulphate was added, and then the precipitate was discarded. The supernatant was applied to a phenyl-Toyopearl 650M column (2.5 cm × 20 cm; Tosoh, Tokyo, Japan) and the aorsin was eluted with a linear gradient of ammonium sulphate at a concentration of 1.4 M to 0 M in buffer A. A pool of active fractions was salted out in 90 % (w/v) saturated ammonium sulphate. After centrifugation, the pellet was solubilized with buffer A containing 50 mM NaCl and was applied to a Sephacryl S-200 HR (Pharmacia, Uppsala, Sweden) column (2.5 cm × 100 cm). Fractions containing aorsin activity were pooled and dialysed against buffer A, and were then applied to a column packed with Arg-Sepharose 4B (1.6 cm × 10 cm; Pharmacia). Elution was performed with a linear gradient of 20–250 mM NaCl. Fractions containing aorsin activity were pooled and dialysed against buffer A, and then applied to a Resource Q column (1 ml; Pharmacia) connected to a fast protein liquid chromatography system (Pharmacia). The column was allowed to re-equilibrate in buffer A before a gradient of 0–10 % buffer B (buffer A containing 1 M NaCl) was applied. Fractions containing active aorsin were pooled and designated as purified aorsin. The purified enzyme migrated as a single band when subjected to SDS/PAGE (10 % gel) performed by the method of Laemmli [17]. Protein concentrations were measured by the method of Bradford [18] with BSA as the standard protein.

Deglycosylation by N-glycosidase F

Freeze-dried aorsin (20 μg) was dissolved in 50 μl of 0.2 M sodium phosphate buffer (pH 8.6) containing 0.5 % (w/v) SDS and 1 % (v/v) 2-mercaptoethanol. Heat denaturation of the preparation was performed at 100 °C for 3 min. An equal volume of 0.2 M sodium phosphate buffer (pH 8.6) containing 3 % (v/v) Nonidet P40 was added before *Flavobacterium* N-glycosidase F was added. After deglycosylation at 37 °C for 18 h, proteins were detected by SDS/PAGE (10 % gel).

Physicochemical characterization

The isoelectric point of the aorsin was determined by isoelectric focusing using a Phastsystem (Pharmacia). The calcium concentration was measured with a Shimadzu AA-660 atomic absorption spectrophotometer (Shimadzu, Kyoto, Japan), which was calibrated with standard solutions of calcium from 0 to 20 μM. Approx. 0.5 mg/ml enzyme was used for determination of the calcium content. The melting temperature (T_m) was determined using a differential scanning calorimeter SSC-560U (Seiko, Kyoto, Japan).

Activation of plasminogen and trypsinogen

Commercial plasminogen was dissolved in 50 mM phosphate buffer, pH 7.5 (buffer C) and applied to a Lys-Sepharose 4B column (1.0 cm × 10 cm; Pharmacia). Elution was performed with 0.2 M of ϵ-amino-n-capronic acid, and the elution fraction containing plasminogen was dialysed against 10 mM Mes buffer, pH 5.4. To activate the plasminogen, 0.6 μg of aorsin was added to 450 μl of purified plasminogen solution, containing 0.5 mg plasminogen, and the mixture was incubated at 37 °C. The enzyme reaction was stopped by adding an equal volume of 0.2 M Tris/HCl buffer, pH 7.5. The generated plasmin hydrolysed a fluorogenic peptide substrate, t-butoxycarbonyl (Boc)-Val-Leu-Lys-MCA. To activate trypsinogen, 1 mg of commercial trypsinogen was dissolved in 1 ml of 10 mM acetate buffer (pH 5.5), 2 μg of aorsin was added, and the mixture was incubated at 30 °C for 12 h. A 10 μl aliquot was mixed with 50 mM

Tris/HCl buffer, pH 7.5, and the trypsin activity was assayed using the substrate benzyloxycarbonyl (Z)-Phe-Arg-MCA. To confirm the accurate processing of the trypsinogen, an aliquot was separated by SDS/PAGE (15% gel) and electroblotted on to a PVDF membrane. The protein fraction was then sequenced.

Amino acid sequence determination

The N-terminal amino acid sequence of purified aorsin was determined by automated sequential Edman degradation using the protein sequencer PPSQ-23 (Shimadzu). To determine the internal amino acid sequence of aorsin, the enzyme was pyridylethylated according to the method of Friedman et al. [19]. The pyridylethylated aorsin was digested with lysyl endopeptidase and the resulting peptides were separated and collected using reverse-phase HPLC. The amino acid sequence of each peptide was sequenced using the above sequencer.

Gene cloning

An *A. oryzae* genomic DNA library was constructed as described previously in [20]. On the basis of the amino acid sequences, deoxyoligonucleotide mixtures were synthesized as follows: sense primer AO-N, 5′-AAYACIGAYCARYTIATHACNCC-3′; antisense primer AO-C1, 5′-RTARTAIGGRTAIGGNGGRTT-3′ (where H indicates A, C or T; I indicates inosine; N indicates A, C, G or T; R indicates A or G and Y indicates C or T). An amplified DNA fragment was obtained by PCR using *A. oryzae* genomic DNA and the above two primers. Using the PCR product as a probe, plaque hybridization was performed and a positive clone was selected. A 4 kb *Xba*I DNA fragment was subcloned into the pBluescript vector and the nucleotide sequence was determined. When the *A. oryzae* niaD300 mutant (*niaD*$^-$) was transformed with this plasmid, transformants produced aorsin activities in wheat-bran solid culture, but did not produce them in submerged culture. Total RNA was purified from the aorsin-producing transformant and a cDNA was amplified by reverse transcriptase-PCR. Consequently, a 2.2 kb DNA fragment was amplified and cloned into *Spe*I and *Cla*I sites of pBluescript and then sequenced.

Expression

An *Spe*I site was generated 14 nucleotides upstream from the start codon of the *aorO* gene encoding aorsin. A 2.8 kb *Spe*I/*Xba*I fragment containing a whole structural gene was inserted into a high-level expression vector plasmid pNAN8142 [21] and *A. oryzae* niaD300 was transformed with the plasmid. The conidia of transformants were selected on a Czapec-Dox plate. The conidia were inoculated in DPYM medium and cultured at 30 °C for 80 h. Aorsin in the culture filtrates was salted out by adding ammonium sulphate up to a 90% (w/v) saturated concentration, and was then purified.

Site-directed mutagenesis

Site-directed mutagenesis of the *aorO* gene was performed by the Kunkel method [22]. The following mutagenic primers were used:

AorE86Q, 5′-AGGCGGGTGGGCAGTCAGATCTC-3′;
AorE86D, 5′-CGGGTGGGGACTCAGATCTCGA-3′;
AorD90N, 5′-AGTCAGATCTCAATTTCGAACTGG-3′;
AorD90E, 5′-TCAGATCTCGAATTCGAACTGGC-3′;
AorD211N, 5′-TTGCTTCTGGTAATAACGGTGTTG-3′;
AorD211E, 5′-GCTTCTGGTGAAAACGGTGTTGC-3′;
AorS354T, 5′-GGTGGAACAACTGCTAGTATGTT-3′;
AorD395N, 5′-CGTCCTGAATAATATCACCAATG-3′;
AorD395E, 5′-GTCCTGAATGAAATCACCAATGG-3′.

All mutations were confirmed by nucleotide sequencing before inserting them into the expression vector pNAN8142.

RESULTS

Purification and properties of aorsin

An enzyme that hydrolyses Z-Arg-Arg-MCA at acidic pH was found in commercial protease M powder that was drained from a solid culture filtrate of *A. oryzae*. We called the enzyme aorsin. The specific activity of the purified aorsin was 12.8 mkat/kg of protein for Z-Arg-Arg-MCA at 30 °C and pH 4.0. Aorsin also hydrolysed clupeine and salmine as well as some fluorogenic peptide substrates, Boc-Gln-Arg-Arg-MCA and Boc-Leu-Lys-Arg-MCA, at an optimum pH of 4.0. Some enzymic and physicochemical properties are summarized in Table 1. The molecular mass of the enzyme was estimated to be 90 kDa and 70–120 kDa by gel filtration and by SDS/PAGE (10% gel) respectively (Figure 1). Protein that was deglycosylated by *Flavobacterium* N-glycosidase F migrated as a sharp single 61 kDa band using SDS/PAGE (10% gel). The results indicated that the enzyme is a monomer with hetero-N-glycosides. The enzyme lost activity irreversibly at pH values of > 6.5 and < 3 during incubation at 30 °C. The half-life ($t_{1/2}$) values for activity loss at 30 °C were 150 min at pH 6.6, 5.5 min at pH 6.8 and 14 min at pH 2.4. Although the enzyme that was treated at 50 °C and pH 4.0 for 10 min remained fully active, the residual activity of the enzyme treated at 55 °C for 10 min was 65%. The $t_{1/2}$ value for loss of activity at 60 °C and pH 4.0 was 3.5 min. The isoelectric point was judged to be 3–4 due to diffuse migration. Aorsin contained 1 g-atom of calcium per mol. The T_m was 66.7 °C. To investigate the inhibition of aorsin, the enzyme was treated with various inhibitors at 30 °C and pH 4.0 for 30 min and the remaining activity was measured. The enzyme activity was inhibited neither by pepstatin nor by the standard inhibitors of peptidases of various types, such as diazoacetyl-DL-norleucine methyl ester (0.05 mM), E-64 (0.1 mM), iodoacetamide (1 mM), PMSF (10 mM), pefabloc SC (10 mM), EDTA (5 mM), EGTA (5 mM) or *o*-phenanthroline (2 mM), at the final concentrations indicated. This inhibition pattern resembled that of pseudomonapepsin (or PSCP), a pepstatin-insensitive carboxylproteinase. Only antipain and leupeptin completely inhibited the aorsin at

Table 1 Some enzymic and physicochemical properties of aorsin

Property	Method, symbol or condition	Value
Molecular mass	Gel filtration	90 kDa
	SDS/PAGE	70–120 kDa
	SDS/PAGE (deglycosylated)	61 kDa
	Deduced from nucleotide sequence	46 522 Da
Isoelectric point	pI	3–4
Ca^{2+} content	Atomic absorption spectrophotometer	1 g atom/mol
Melting temperature (T_m)	Differential scanning calorimeter	66.7 °C
Optimum pH	pH	4.0
Specific activity	Z-Arg-Arg-MCA, 30 °C, pH 4.0	12.8 mkat/kg
Half-life ($t_{1/2}$) for tolerance of pH	30 °C, pH 6.6	150 min
	30 °C, pH 6.8	5.5 min
	30 °C, pH 2.4	14 min
Half-life ($t_{1/2}$) for tolerance of temperature	60 °C, pH 4.0	3.5 min
Inhibitor	Antipain, leupeptin	

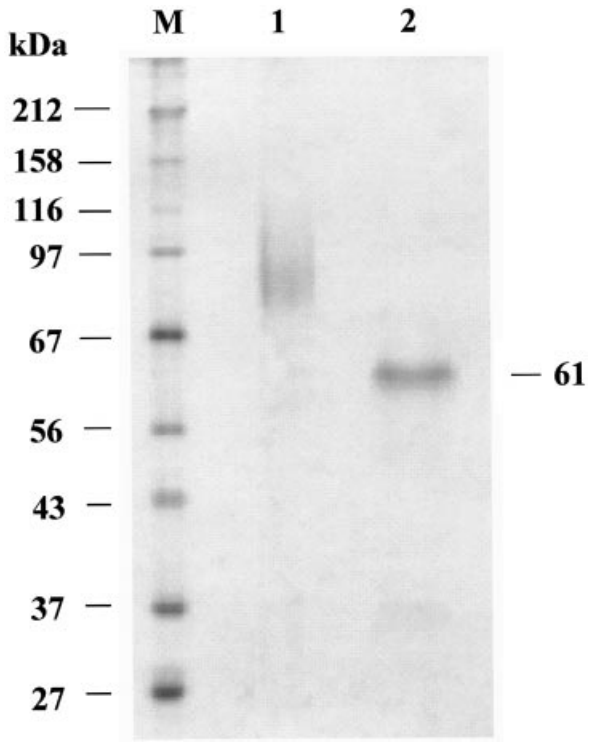

Figure 1 Analysis of aorsin from *A. oryzae* by SDS/PAGE

Approx. 2 μg of purified aorsin were separated under reducing-gel conditions (10% gel) and stained with Coomassie Brilliant Blue. M represents the molecular mass marker. Note the appearance of a diffuse staining band at 70–120 kDa (lane 1), and the shift in mobility of aorsin that was deglycosylated with N-glycosidase F (lane 2).

0.01 mM, but aorsin was not inhibited by tyrostatin [23], which is a specific inhibitor of PSCP.

Substrate specificity

To investigate the substrate specificity of aorsin, 45 types of fluorogenic peptide substrates were used. Aorsin hydrolysed the substrates containing the basic amino acid residues arginine or lysine at the P_1 position, but did not hydrolyse any other substrates possessing non-basic amino acids at the P_1 position, indicating that the enzyme has trypsin-like activity. Several good fluorogenic substrates and their kinetic parameters are shown in Table 2. Although no significant difference in k_{cat}/K_m values was observed between the substrates, relatively low k_{cat}/K_m values and high K_m values were observed for Boc-Gln-Gly-Arg-MCA and Boc-Leu-Gly-Arg-MCA in which the P_2 positions are occupied by glycine. Low K_m values were also observed for substrates having dibasic amino acids at the P_1 and P_2 positions, such as Boc-Gln-Arg-Arg-MCA, Boc-Leu-Arg-Arg-MCA and Boc-Gly-Arg-Arg-MCA. The highest k_{cat}/K_m value was observed for Boc-Leu-Lys-Arg-MCA. On the other hand, aorsin hydrolysed a peptide hormone dynorphin A (Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Ile-Arg-Pro-Lys-Leu-Lys) at the peptide bond between Arg^7 and Ile^8, but other peptide bonds were not cleaved despite the presence of several basic residues. These findings indicate that aorsin has a unique, specific requirement for a basic residue at P_1 and to prefer paired basic residues. In terms of cleavage, aorsin is closer to a processing enzyme than it is to trypsin.

Plasminogen that was treated with aorsin hydrolysed a Boc-Val-Leu-Lys-MCA substrate, which is a substrate of plasmin,

Table 2 Kinetic parameters of aorsin towards fluorogenic peptide substrates at pH 4.0

The parameters were determined as described in the Experimental section. The k_{cat} value was calculated using molecular mass 46522 Da. The values of k_{cat}, K_m and k_{cat}/K_m are the means $\pm$ S.D. for three to five independent experiments.

Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
Boc-Gln-Arg-Arg-MCA	0.7 ± 0.03	0.10 ± 0.01	$(1.45 \pm 0.06) \times 10^5$
Boc-Leu-Arg-Arg-MCA	3.9 ± 0.25	0.24 ± 0.02	$(6.12 \pm 0.45) \times 10^4$
Boc-Gly-Arg-Arg-MCA	2.8 ± 0.22	0.25 ± 0.02	$(9.01 \pm 0.47) \times 10^4$
Boc-Leu-Lys-Arg-MCA	10.0 ± 0.51	5.80 ± 0.37	$(5.79 \pm 0.38) \times 10^5$
Boc-Gly-Lys-Arg-MCA	26.8 ± 1.33	2.69 ± 0.18	$(1.00 \pm 0.07) \times 10^5$
Boc-Leu-Thr-Arg-MCA	58.0 ± 2.47	4.11 ± 0.12	$(7.09 \pm 0.38) \times 10^4$
Boc-Leu-Gly-Arg-MCA	84.5 ± 3.75	1.26 ± 0.04	$(1.49 \pm 0.09) \times 10^4$
Boc-Gln-Gly-Arg-MCA	88.3 ± 4.06	0.84 ± 0.05	$(9.46 \pm 0.07) \times 10^3$

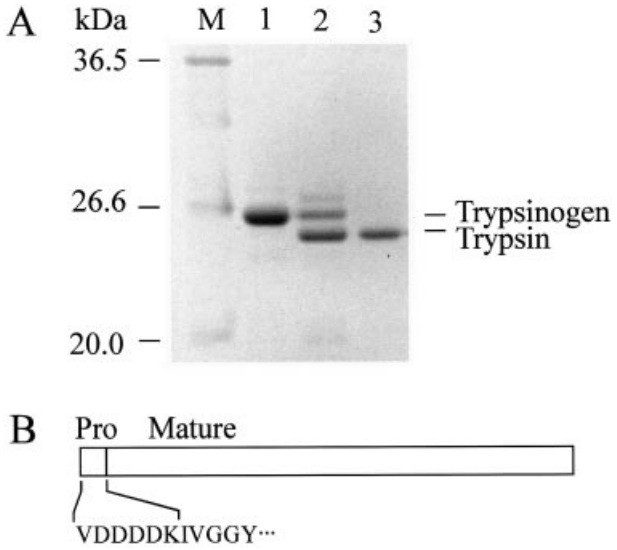

Figure 2 Conversion of trypsinogen into trypsin by aorsin

The activation conditions of trypsinogen and the enzyme assay conditions of converted trypsin are described in the Experimental section. (**A**) The reaction mixtures for the trypsinogen activation were analysed by SDS/PAGE on a 15% gel and stained with Coomassie Brilliant Blue R-250. Trypsinogen was stained with apparent homogeneity (lane 1). The reaction product, converted into trypsin, migrated slightly faster than trypsinogen (lane 2), and this migration was similar to that of authentic trypsin (lane 3). (**B**) The natural processing site of bovine trypsinogen (24 kDa). The N-terminal sequence of the converted trypsin (lane 2 in **A**) was determined to be IVGGY … .

whereas aorsin or plasminogen alone did not hydrolyse the substrate. However, we failed to determine the processing site because activated plasmin could not be purified, probably due to further degradation. Trypsinogen was also activated by being treated with aorsin, and the generated trypsin hydrolysed Z-Phe-Arg-MCA, which was not cleaved by aorsin. The N-terminal sequence of the activated trypsin was determined to be Ile-Val-Gly-Gly-Tyr-, indicating that accurate processing had occurred by aorsin as shown in Figure 2.

Analysis of primary structure

A gene (*aorO*) was cloned from the *A. oryzae* genomic library. To confirm whether the gene is functional, a plasmid containing a 4 kb *Xba*I fragment was used to transform the *A. oryzae* niaD300

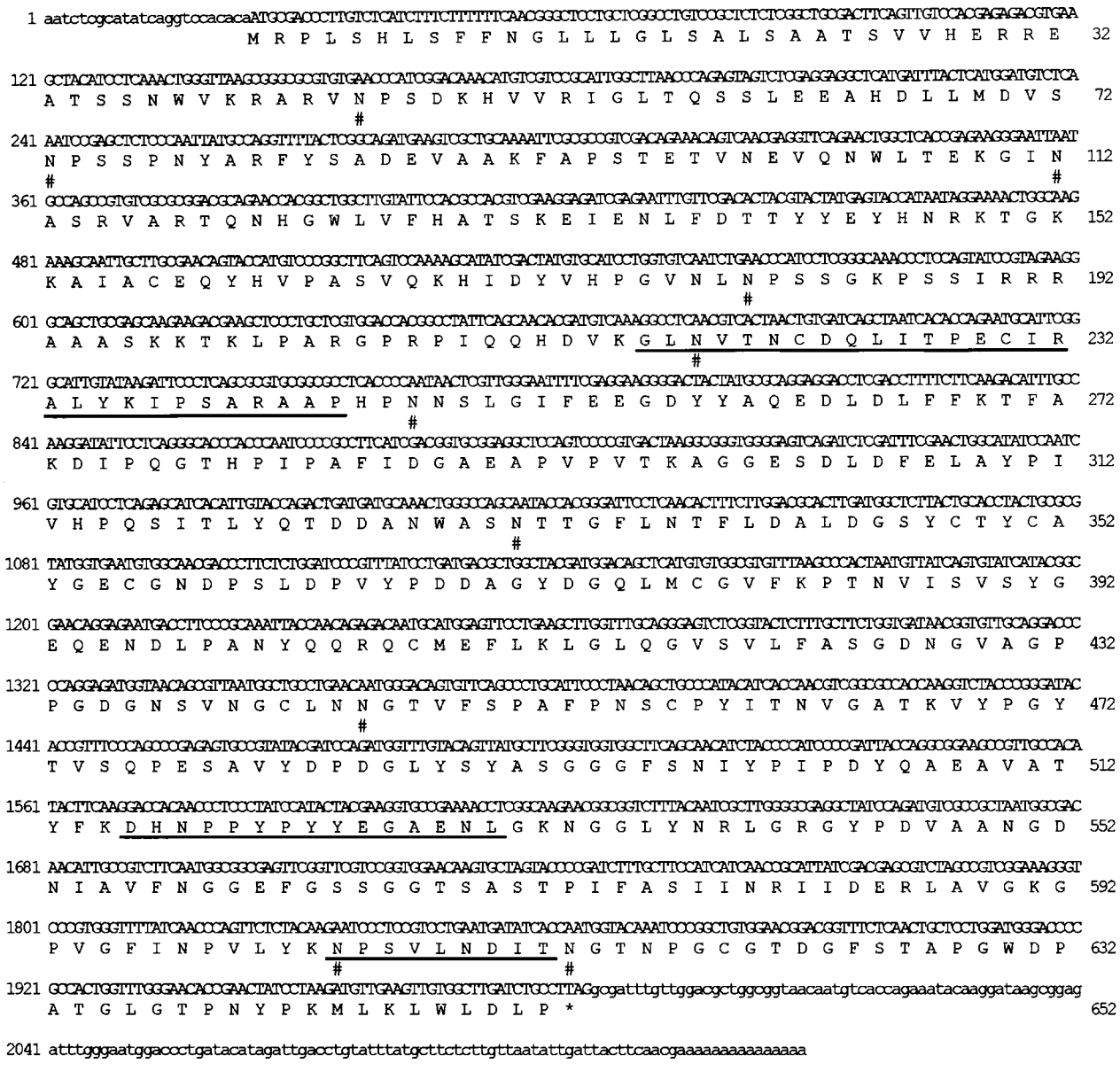

Figure 3 Nucleotide and deduced amino acid sequences of the *aorO* gene

Polyadenylation was observed at the 3′ ends of the cloned cDNA. Untranslated regions are shown in lower case letters. The N-terminal amino acid sequence of the purified aorsin and the internal amino acid sequences isolated after lysyl endopeptidase digestion are underlined. The # symbols indicate the positions of potentially N-linked glycosylation. An asterisk (*) indicates the stop codon. The nucleotide sequence data appear in the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the Accession Number AB084899.

mutant strain. Consequently the enzyme activity of the transformant was observed to be about 10-fold higher than that of the recipient strain in the solid culture, but neither demonstrated any activity in the submerged culture. This finding indicates that the cloned gene is functional and expressed only in the solid culture state.

A 4 kb *Xba*I fragment of the cloned gene and a cDNA were sequenced and compared. Eight introns with short sequences of 50–75 nucleotides were found in the gene. The complementary DNA sequence is shown in Figure 3. The complete structural gene encoded a protein of 652 amino acids. The amino acid sequences of the lysyl-endopeptidase-digested peptides were verified at positions Asp^{516} to Leu^{530} and Asn^{604} to Thr^{612}. The N-terminal amino acid sequence of the mature enzyme was found at positions Gly^{216} to Pro^{244}. Consequently the prepropeptide consisted of 215 amino acid residues from Met^{1} to Lys^{215}, and the mature form of the enzyme was composed of 437 amino acid residues. The deduced amino acid sequence of the mature protein contained six potential sites for asparagine-linked glycosylation at positions 218, 247, 331, 445, 604 and 613. A protein identity search revealed that the mature aorsin protein was 35% identical with CLN2p [8], 23% identical with PSCP [6] and 22% identical with *Xanthomonas* pepstatin-insensitive carboxyl proteinase (XCP) [13], all of which are pepstatin-insensitive carboxyl proteinases (Figure 4). Several significant amino acids were conserved in aorsin, at Glu^{86}, Asp^{90}, Asp^{211}, Ser^{354} and

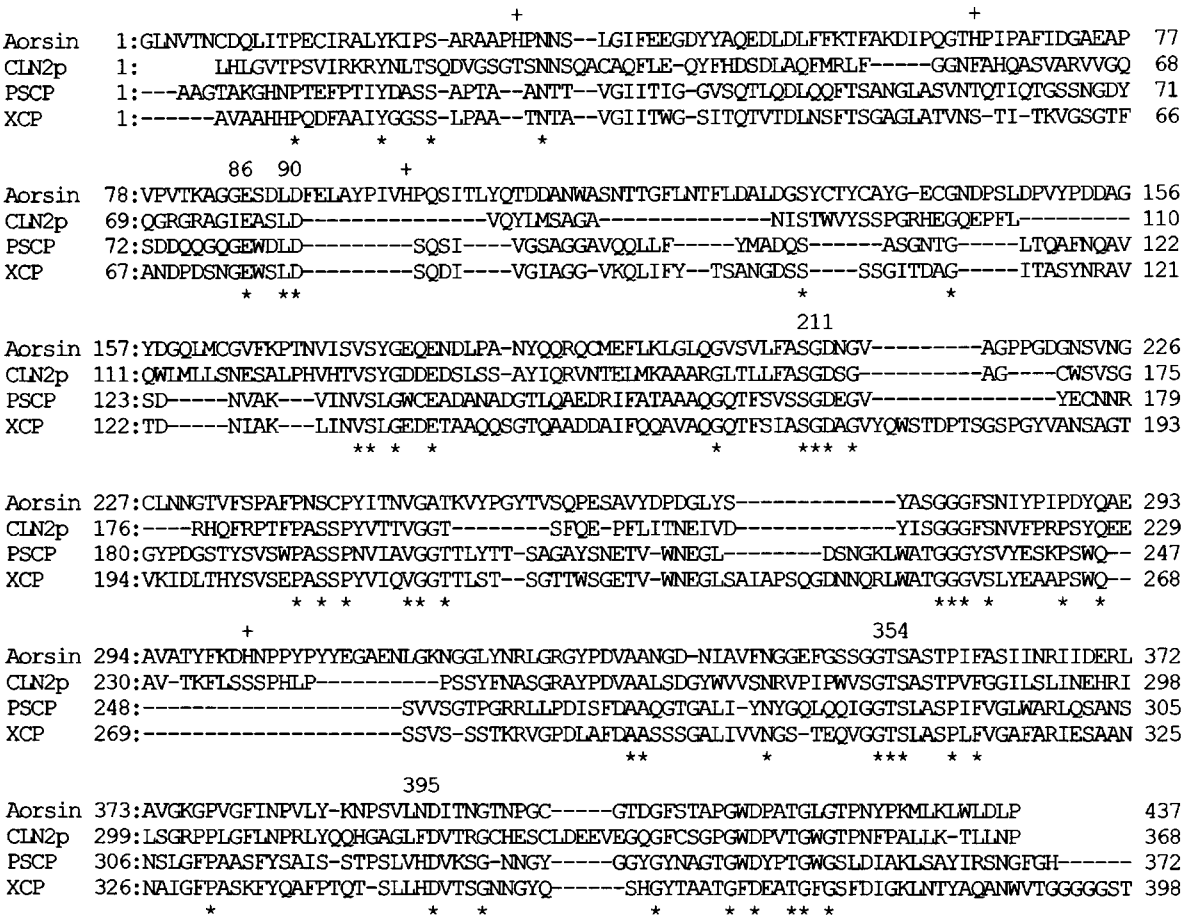

Figure 4 Comparison of the amino acid sequences of pepstatin-insensitive carboxyl proteinases (serine-carboxyl proteinases)

The amino acid sequence of aorsin is compared with those of CLN2p [8], PSCP [6] and XCP [13]. Gaps (—) were introduced to obtain maximum alignment. The numbering above the sequences is that of aorsin. Identical amino acid residues between aorsin and other pepstatin-insensitive carboxyl proteinases are indicated by asterisks (*). Four histidine residues in aorsin are indicated by + symbols.

Table 3 Kinetic parameters for hydrolysis of a fluorogenic peptide substrate Boc-Leu-Lys-Arg-MCA catalysed by aorsin and its mutants at pH 4.0

The parameters were determined as described in the Experimental section and Table 2.

Enzyme	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}\,s^{-1}$)	Relative k_{cat}/K_m
Wild-type	10.0 ± 0.51	(5.80 ± 0.37)	$(5.79 \pm 0.38) \times 10^5$	1
$Glu^{86} \rightarrow Asp$	63.6 ± 4.80	$(3.74 \pm 0.22) \times 10^{-2}$	$(5.89 \pm 0.27) \times 10^2$	1×10^{-3}
$Glu^{86} \rightarrow Gln$	19.7 ± 0.23	$(1.55 \pm 0.02) \times 10^{-3}$	$(7.87 \pm 0.34) \times 10$	1.4×10^{-4}
$Asp^{90} \rightarrow Glu$	117.0 ± 5.85	$(3.52 \pm 0.14) \times 10^{-2}$	$(3.00 \pm 0.17) \times 10^2$	5.2×10^{-4}
$Asp^{90} \rightarrow Asn$	22.5 ± 1.87	$(6.15 \pm 0.31) \times 10^{-1}$	$(2.75 \pm 0.19) \times 10^4$	0.048
$Asp^{211} \rightarrow Glu$	23.6 ± 1.51	$(3.23 \pm 0.21) \times 10^{-4}$	$(1.37 \pm 0.05) \times 10$	2.4×10^{-5}
$Asp^{211} \rightarrow Asn$	16.1 ± 1.38	$(1.19 \pm 0.05) \times 10^{-3}$	$(7.41 \pm 0.42) \times 10$	1.3×10^{-4}
$Ser^{354} \rightarrow Thr$	19.3 ± 1.23	$(6.45 \pm 0.11) \times 10^{-4}$	$(3.34 \pm 0.25) \times 10$	5.8×10^{-5}
$Asp^{395} \rightarrow Glu$	10.6 ± 0.92	(5.26 ± 0.26)	$(4.97 \pm 0.33) \times 10^5$	0.86
$Asp^{395} \rightarrow Asn$	9.6 ± 0.62	(4.49 ± 0.09)	$(4.68 \pm 0.26) \times 10^5$	0.81

Asp^{395}, each of which was equivalent to Glu^{80}, Asp^{84}, Asp^{170}, Ser^{287} and Asp^{328} in PSCP respectively, all of which have previously been described as residues essential for enzyme activity [15,24].

Analysis of mutant enzymes

To identify the catalytic residues, the *aorO* gene was mutated at the site of encoding Glu^{86}, Asp^{90}, Asp^{211}, Ser^{354} or Asp^{395}. Since we confirmed that an authentic *aorO* gene product, aorsin, was not detected in a submerged culture broth of *A. oryzae*, we employed *A. oryzae* as a host for recombinant aorsin production. Mutated enzymes were secreted into submerged culture broths in mature form. Purified recombinant enzymes showed that they have the same patterns as the authentic aorsin in terms of molecular mass, carbohydrate moiety and N-terminal sequence.

The results of site-directed mutagenesis of aorsin are shown in Table 3. An aspartic acid residue, Asp^{328} in PSCP and Asp^{517} in CLN2p, has been identified as a catalytic residue [14,24] or a ligand of calcium binding [15]. However, mutant aorsins $Asp^{395} \rightarrow Asn$ and $Asp^{395} \rightarrow Glu$ had full activity and lost their calcium-binding ability completely (calcium was not detected by atomic absorption spectrophotometry). This finding indicates that Asp^{395} is a calcium-binding site and is not related to the catalytic function. A mutant enzyme replacing Asp^{211}, equivalent to Asp^{170} in PSCP and Asp^{360} in CLN2p, with either asparagine ($Asp^{211} \rightarrow Asn$) or glutamic acid ($Asp^{211} \rightarrow Glu$) showed that the k_{cat}/K_m values were remarkably reduced by 7800–42000-fold. Mutation of Ser^{354}, which is equivalent to Ser^{287} in PSCP and

Ser^{475} in CLN2p, with threonine (Ser^{354} → Thr) reduced the k_{cat}/K_m value by 17000-fold. Substitution of Glu^{86}, equivalent to Glu^{80} in PSCP, with either aspartic acid (Glu^{86} → Asp) or glutamine (Glu^{86} → Gln) also reduced the k_{cat}/K_m value by 1000–7400-fold. Surprisingly, however, the k_{cat}/K_m value of the replacement of Asp^{90}, equivalent to Asp^{84} in PSCP, with asparagine (Asp^{90} → Asn) was only 21-fold lower than the k_{cat}/K_m value of the wild-type enzyme.

DISCUSSION

Trypsin-like activity was found in *A. oryzae* using a fluorogenic peptide substrate, Z-Arg-Arg-MCA. Although this activity was not detected in the submerged culture broth, at least two enzymes were detected in the solid culture filtrates of wheat bran or rice, and they were active at nearly pH 4.0 and pH 9.0. The enzyme that was active at alkaline pH was a deuterolysin, which has been previously described in [25] and the other enzyme, which is active at acidic pH, is a novel acid proteinase, which we called aorsin. Aorsin activity increased at the initial growth stage in solid culture and decreased in the late growth stage. These results suggest that aorsin is a solid-culture-specific protein, but it is not essential for growth.

The purified enzyme migrated diffusely with a molecular mass of 70–120 kDa using SDS/PAGE (10 % gel), although the molecular mass predicted from the sequence of the *aorO* gene was only 46.5 kDa. The high apparent molecular mass and the presence of six potential asparagine-linked glycosylation sites in the aorsin sequence suggested that aorsin is highly glycosylated. Digestion of aorsin with N-glycosidase F, which removes asparagine-linked oligosaccharides, resulted in an increased mobility estimated to be 61 kDa by SDS/PAGE (10 % gel), but not the 46.5 kDa predicted from the *aorO* gene, suggesting that aorsin also contains serine/threonine-linked oligosaccharides.

Substrate specificities of acid proteinases have been well characterized. Most of them prefer hydrophobic or bulky amino acid residues at the P_1 position of the substrate, and only glycosylphosphatidylinositol-anchored aspartic proteinase yapsins are basic-specific proteinases [26,27]. Aorsin is similar to yapsin in specificity and acid preference, but it is a cell-free serine proteinase. Summarizing the analysis of our specificity studies, we conclude that aorsin appears to require a basic residue at P_1, and to prefer paired basic residues.

Most acid proteinases have been referred to as aspartic proteinases, but recently the crystal structure of PSCP active at pH 3.5 has been determined and its unique catalytic triad was defined as Glu^{80}, Asp^{84} and Ser^{287}. This has led to the enzyme being considered as a serine proteinase [15,28]. The fold in PSCP corresponds to that of subtilisin, and Glu^{80} is a structural equivalent to His^{64} in subtilisin. More recently, the 1.4 Å crystal structure (where 1 Å = 0.1 nm) of kumamolysin serine-carboxyl proteinase (KSCP) has been solved [29]. The structural insight into KSCP is in close agreement with that of PSCP. The PSCP and KSCP have almost identical catalytic centres. The crystal structures and the sequence similarity suggest that Glu^{86}, Asp^{90} and Ser^{354} of aorsin are the catalytic residues and that Asp^{211} is involved in forming the oxyanion hole and Asp^{395} is a calcium ligand. In the present study, we have performed mutational analyses. The k_{cat}/K_m values of the mutant enzymes Glu^{86} → Gln, Asp^{211} → Asn and Ser^{354} → Thr were 3–4 orders of magnitude lower than that of wild-type aorsin, while Asp^{90} → Asn was only 21-fold lower. We therefore conclude that Ser^{354} and Glu^{86} are considered important, but the role of Asp^{90} is more minor. Our mutational experiments suggest that the roles of the residues mentioned above are almost identical with those of PSCP and KSCP. However, to explain why the Asp^{90} mutants have sufficient activity, further studies are necessary, including X-ray crystallographic analysis of the three-dimensional structure.

Aorsin is the first identified serine proteinase with trypsin-like specificity at acidic pH.

REFERENCES

1 Barrett, A. J., Rawlings, N. D. and Woessner, J. F. (eds.) (1998) Aspartic peptidases. In Handbook of Proteolytic Enzymes, pp. 799–986, Academic Press, San Diego, CA

2 Umezawa, H., Aoyagi, T., Morishima, H., Matsuzaki, M., Hamada, M. and Takeuchi, T. (1970) Pepstatin, a new pepsin inhibitor produced by actinomycetes. J. Antibiot. (Tokyo) **23**, 259–262

3 Takahashi, K. (1995) Proteinase A from *Aspergillus niger*. Methods Enzymol. **248**, 146–155

4 Murao, S., Oda, K. and Matsushita, Y. (1972) New acid protease from *Scytalidium lignicolum*. Agric. Biol. Chem. **36**, 1647–1650

5 Oda, K., Sugitani, M., Fukuhara, K. and Murao, S. (1987) Purification and properties of a pepstatin-insensitive carboxyl proteinase from a Gram-negative bacterium. Biochim. Biophys. Acta **923**, 463–469

6 Oda, K., Takahashi, T., Tokuda, Y., Shibano, Y. and Takahashi, S. (1994) Cloning, nucleotide sequence, and expression of an isovaleryl pepstatin-insensitive carboxyl proteinase gene from *Pseudomonas* sp. 101. J. Biol. Chem. **269**, 26518–26524

7 Oda, K., Nakazima, T., Terashita, T., Suzuki, K. and Murao, S. (1987) Purification and properties of an S-PI (pepstatin Ac)-insensitive carboxyl proteinase from *Xanthomonas* sp. bacterium. Agric. Biol. Chem. **51**, 3073–3080

8 Sleat, D. E., Donnelly, R. J., Lackland, H., Liu, C.-G., Sohar, I., Pullarkat, R. K. and Lobel, P. (1997) Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. Science **277**, 1802–1805

9 Wisniewski, K. E., Kaczmarski, A., Kida, E., Connell, F., Kaczmarski, W., Michaelewski, M. P., Moroziewicz, D. N. and Zhong, N. (1999) Reevaluation of neuronal ceroid lipofuscinoses: atypical juvenile onset may be the result of CLN2 mutations. Mol. Genet. Metab. **66**, 248–252

10 Junaid, M. A., Wu, G. and Pullarkat, R. K. (2000) Purification and characterization of bovine brain lysosomal pepstatin-insensitive proteinase, the gene product deficient in the human late-infantile neuronal ceroid lipofuscinosis. J. Neurochem. **74**, 287–294

11 Rawlings, N. D. and Barrett, A. J. (1999) Tripeptidyl-peptidase I is apparently the CLN2 protein absent in classical late-infantile neuronal ceroid lipofuscinosis. Biochim. Biophys. Acta **1429**, 496–500

12 Ezaki, J., Takeda-Ezaki, M., Oda, K. and Kominami, E. (2000) Characterization of endopeptidase activity of tripeptidyl peptidase-I/CLN2 protein which is deficient in classical late infantile neuronal ceroid lipofuscinosis. Biochem. Biophys. Res. Commun. **268**, 904–908

13 Oda, K., Ito, M., Uchida, K., Shibano, Y., Fukuhara, K. and Takahashi, S. (1996) Cloning and expression of an isovaleryl pepstatin-insensitive carboxyl proteinase gene from *Xanthomonas* sp. T-22. J. Biochem. (Tokyo) **120**, 564–572

14 Lin, L., Sohar, I., Lackland, H. and Lobel, P. (2001) The human CLN2 protein/trypeptidyl-peptidase I is a serine protease that autoactivates at acidic pH. J. Biol. Chem. **276**, 2249–2255

15 Wlodawer, A., Li, M., Dauter, Z., Gustchina, A., Uchida, K., Oyama, H., Dunn, B. M. and Oda, K. (2001) Carboxyl proteinase from *Pseudomonas* defines a novel family of subtilisin-like enzymes. Nat. Struct. Biol. **8**, 442–446

15a Rawlings, N. D., O'Brien, E. A. and Barrett, A. J. (2002) MEROPS: the protease database. Nucleic Acids Res. **30**, 343–346

16 Minetoki, T., Nunokawa, Y., Gomi, K., Kitamoto, K., Kumagai, C. and Tamura, G. (1996) Deletion analysis of promoter elements of the *Aspergillus oryzae agdA* gene encoding α-glucosidase. Curr. Genet. **30**, 432–438

17 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**, 680–685

18 Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–254

19 Friedman, M., Krull, L. H. and Cavins, J. F. (1970) The chromatographic determination of cystine and cysteine residues in proteins as S-β-(4-pyridylethyl)cysteine. J. Biol. Chem. **245**, 3868–3871

20 Lee, B. R., Kitamoto, K., Yamada, O. and Kumagai, C. (1995) Cloning, characterization and overproduction of nuclease S1 gene (*nucS*) from *Aspergillus oryzae*. Appl. Microbiol. Biotechnol. **44**, 425–431

21 Ozeki, K., Kanda, A., Hamachi, M. and Nunokawa, Y. (1996) Construction of a promoter probe vector autonomously maintained in *Aspergillus* and characterization of promoter regions derived from *A. niger* and *A. oryzae* genomes. Biosci. Biotechnol. Biochem. **60**, 383–389
22 Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. U.S.A. **82**, 488–492
23 Oda, K., Fukuda, Y., Murao, S., Uchida, K. and Kainosho, M. (1989) A novel proteinase inhibitor, tyrostatin, inhibiting some pepstatin-insensitive carboxyl proteinases. Agric. Biol. Chem. **53**, 405–415
24 Oyama, H., Abe, S., Ushiyama, S., Takahashi, S. and Oda, K. (1999) Identification of catalytic residues of pepstatin-insensitive carboxyl proteinases from prokaryotes by site-directed mutagenesis. J. Biol. Chem. **274**, 27815–27822
25 Fushimi, N., Ee, C. E., Nakajima, T. and Ichishima, E. (1999) Aspzincin, a family of metalloendopeptidases with a new zinc-binding motif. J. Biol. Chem. **274**, 24195–24201
26 Olsen, V., Guruprasad, K., Cawley, N. X., Chen, H.-C., Blundell, T. M. and Loh, Y. P. (1998) Cleavage efficiency of the novel aspartic protease yapsin 1 (Yap3p) enhanced for substrates with arginine residues flanking the P1 site: correlation with electronegative active-site pockets predicted by molecular modeling. Biochemistry **37**, 2768–2777
27 Komano, H., Rockwell, N., Wang, G. T., Krafft, G. A. and Fuller, R. S. (1999) Purification and characterization of the yeast glycosylphosphatidylinositol-anchored, monobasic-specific aspartyl protease yapsin 2 (Mkc7p). J. Biol. Chem. **274**, 24431–24437
28 Wlodawer, A., Li, M., Gustchina, A., Dauter, Z., Uchida, K., Oyama, H., Goldfarb, N. E., Dunn, B. M. and Oda, K. (2001) Inhibitor complexes of the *Pseudomonas* serine-carboxyl proteinase. Biochemistry **40**, 15602–15611
29 Comellas-Bigler, M., Fuentes-Prior, P., Maskos, K., Huber, R., Oyama, H., Uchida, K., Dunn, B. M., Oda, K. and Bode, W. (2002) The 1.4 Å crystal structure of kumamolysin: a thermostable serine-carboxyl-type proteinase. Structure **10**, 865–876

Received 30 October 2002/18 December 2002; accepted 8 January 2003
Published as BJ Immediate Publication 8 January 2003, DOI 10.1042/BJ20021691